

Robot futuri Domenico Parisi

1. Vita artificiale

La vita artificiale è lo studio dei fenomeni del mondo vivente attraverso la loro riproduzione in sistemi artificiali, simulati con il computer o fisicamente realizzati. La vita artificiale ha una visione ampia e integrata dei fenomeni del mondo vivente, dove per “ampia” si intende che fa rientrare tra i fenomeni del mondo vivente non solo i fenomeni studiati dalle scienze biologiche ma anche quelli studiati dalle scienze cognitive e dalle scienze sociali, e per “integrata” che cerca di applicare a tutti questi fenomeni gli stessi modelli e gli stessi schemi interpretativi: le reti neurali, modelli di spiegazione del comportamento ispirati alla struttura e al modo di funzionare del cervello, i modelli ad agente, i modelli di sistemi complessi, quelli evolutivi e quelli a rete. La vita artificiale riconosce nei fenomeni biologici, cognitivi e sociali tre successivi stadi temporali della realtà, con i fenomeni comportamentali e cognitivi emersi dopo i fenomeni biologici e i fenomeni delle società umane emersi dopo l’emergere di Homo sapiens e della sua mente, e spinge a porsi importanti domande generali quali: Perché la realtà è un continuo emergere di nuove classi di fenomeni? Quali principi generali governano l’emergere di fenomeni nuovi? Quali fenomeni nuovi potranno emergere in futuro?

La riproduzione della realtà in sistemi artificiali è una nuova metodologia della scienza, che si affianca a quella tradizionale che consiste nell’osservare la realtà e nello spiegarla con teorie espresse a parole o con le formule matematiche. Le teorie espresse a parole sono spesso mal definite e non è chiaro quali previsioni empiriche possono essere tratte da esse, mentre le teorie espresse con le formule matematiche, pur essendo precise e con chiare previsioni empiriche, non possono essere applicate a tutti i fenomeni. Le simulazioni sono un nuovo modo di esprimere le teorie o i modelli della scienza, traducendo le teorie o i modelli in programmi di calcolatore. Le simulazioni possono essere applicate a qualunque fenomeno, e perciò costituiscono un fondamentale strumento di integrazione tra le discipline, sono necessariamente ben definite perché altrimenti il programma non si può scrivere o non gira nel calcolatore, e fanno

chiare predizioni empiriche in quanto i risultati della simulazione sono le predizioni empiriche della teoria incorporata nella simulazione. Le simulazioni sono anche strumenti per pensare. Se fino ad oggi si poteva dire (e qualcuno l'ha detto) "Cerca le parole e troverai i pensieri", oggi si può dire "Costruisci una simulazione e troverai i pensieri".

Le caratteristiche della vita artificiale indicano anche quali ostacoli essa incontra e incontrerà ad affermarsi. La vita artificiale è essenzialmente non disciplinare mentre la scienza tradizionalmente è divisa in discipline. Esprimendo le sue teorie sotto forma di sistemi artificiali, la vita artificiale necessariamente formula teorie e modelli espliciti e precisi, mentre molte discipline scientifiche sono soprattutto empiriche (ad esempio le scienze biologiche) o non hanno attitudine a formulare teorie precise e quantitative (come spesso le scienze cognitive e sociali). La vita artificiale stabilisce una completa continuità tra le scienze naturali e le scienze dell'uomo, e questo suscita resistenze di natura culturale, religiosa e politica.

In questo lavoro voglio argomentare che, quando viene applicata allo studio del comportamento e della cognizione, la vita artificiale è essenzialmente robotica. Ma i robot di oggi sono povere immagini di quelli che potranno e dovranno essere i robot del futuro. Sono questi robot del futuro che vorrei delineare in questo scritto.

2. I robot sono la nuova icona della mente

Per tutta la seconda metà del Novecento l'icona, il simbolo, della mente è stato il computer. Il computer è una macchina che scambia informazioni con il mondo esterno, conserva le informazioni e le ritrova nella sua memoria, riconosce gli oggetti, usa il linguaggio, ragiona, decide, cioè fa le cose che fa la mente umana. Per questo appena il computer è stato inventato, a metà del Novecento, è stato inevitabile che si vedesse una analogia tra la mente umana e il computer: la mente è un computer. Siccome il computer fa tutte le cose che fa limitandosi a manipolare simboli sulla base di comandi anch'essi espressi con simboli, l'analogia tra la mente e il computer ha costituito la base di una visione della mente come manipolazione di simboli: la teoria computazionale della mente.

L'idea che la mente sia un computer e che quindi le capacità umane si possano analizzare e spiegare come si analizzano e si spiegano le capacità del computer è stata accolta con entusiasmo dalla scienza della mente, la psicologia. La psicologia è da sempre afflitta dal problema che le entità e i processi mentali che essa chiama in causa per spiegare le capacità e i comportamenti tendono ad essere qualcosa di poco preciso e di sfuggente in quanto non sono direttamente osservabile e misurabili. La teoria computazionale della

mente viene in aiuto della psicologia perché dimostra che della mente si può parlare con la stessa precisione e oggettività con cui gli informatici parlano del computer, dato che secondo la teoria computazionale della mente, la mente è una macchina come è una macchina il computer.

Il computer però è una macchina molto particolare: è una macchina senza corpo, senza fisicità. Che un computer sia fatto fisicamente in un certo modo, cioè abbia certe dimensioni, una certa forma, che dentro al computer avvengano certi processi fisici invece che altri, è irrilevante per quello che il computer deve fare. L'unica cosa importante è che il computer sia in grado di operare su dei simboli applicando regole espresse con dei simboli. Il modo in cui il computer funziona è diverso dal modo in cui funziona il cervello. Il computer funziona in modo sequenziale, eseguendo un comando dopo l'altro, quello che si chiama il programma. Il cervello invece funziona in modo parallelo, con tanti processi che avvengono nello stesso momento in diversi punti del cervello e che si influenzano l'uno con l'altro, senza che ci sia nessuno a comandare, senza un programma.

Ma il computer è una macchina senza corpo, senza fisicità, non solo perché il modo in cui funziona non ha nulla a che vedere con il modo in cui funziona il sistema fisico che controlla il comportamento degli organismi, cioè il cervello. Il computer è una macchina senza corpo anche perché non ha sensi (occhi, orecchie) né organi motori (gambe, braccia, mani) con cui poter interagire fisicamente con il mondo esterno. Il computer interagisce soltanto con chi lo usa, e si tratta di una interazione puramente simbolica. Il computer invia simboli a chi lo usa facendoli comparire sullo schermo del computer, e chi lo usa invia simboli al computer scrivendo sulla tastiera o muovendo il mouse. Il computer non ha nessuna autonomia; è una semplice appendice di noi esseri umani. Proprio il contrario non solo di un essere umano ma di qualunque organismo vivente.

La teoria computazionale della mente autorizza a studiare la mente senza considerare la sua fisicità, ma questo significa che i grandi e continui progressi che le scienze biologiche fanno nello studio del cervello e del corpo sono irrilevanti per la scienza della mente e la scienza della mente non può giovare più di tanto di questi progressi. E questo è un peccato. E poi c'è un argomento più generale che va contro la teoria computazionale della mente. L'analogia tra la mente e il computer crea una separazione netta tra le scienze della natura e le scienze della mente. Ma se la scienza è fatta di discipline separate, la realtà non è disciplinare. La realtà è fatta di fenomeni che sono sì diversi ma sono collegati gli uni con gli altri e si spiegano gli uni con gli altri. Non è possibile una scienza della mente che non si

fondi su una separazione tra la scienza della natura e la scienza della mente?

Inoltre sempre più chiaramente dalle ricerche di intelligenza artificiale emerge un fatto che costituisce un imbarazzo per la teoria computazionale della mente: le cose che riescono facili al computer, riescono difficili agli esseri umani, e viceversa, quelle che riescono facili agli esseri umani, sono difficili per il computer. I computer sono bravissimi a trovare le informazioni dalle loro immense memorie e a fare lunghi e complessi ragionamenti senza commettere neppure un errore, mentre per gli esseri umani certamente non è così. Gli esseri umani sono bravissimi a riconoscere gli oggetti e le situazioni in qualunque modo e condizione si presentino, ad usare uno strumento così complicato come linguaggio con grande facilità, a muoversi in ogni tipo di ambiente con grande disinvoltura, a provare sentimenti e emozioni, mentre i computer hanno enormi difficoltà a fare tutte queste cose. Come è possibile una cosa del genere se la mente è un computer?

Allora, verso la fine del Novecento è andata emergendo una nuova scienza della mente che, come la teoria computazionale, interpreta la mente in analogia con un artefatto, una cosa fatta da noi, come è il computer, ma non si tratta più del computer. La nuova scienza della mente elimina ogni separazione tra la scienza della natura e la scienza della mente. Pensa che i comportamenti e le capacità vadano spiegati chiamando in causa nient'altro che il cervello, il resto del corpo degli organismi, e l'ambiente fisico in cui gli organismi stanno e con cui interagiscono. Lo strumento di base di questa nuova scienza della mente sono le reti neurali, modelli semplificati del cervello che spiegano il comportamento chiamando in causa soltanto unità che corrispondono ai neuroni e connessioni che corrispondono alle sinapsi tra neuroni. Dietro alla mente, come dietro ad ogni fenomeno della natura, non ci sono che cause fisiche che producono effetti fisici e tutto è intrinsecamente e fino in fondo quantitativo.

Le reti neurali all'inizio sono rimaste prive di un corpo e prive di un ambiente fisico con cui interagire ma poi sono state messe dentro un corpo, con le sue dimensioni, la sua forma, i suoi sensori e i suoi organi motori, e il corpo è stato messo dentro a un ambiente fisico. Il comportamento e le capacità hanno la loro base e la loro origine nelle interazioni fisiche del corpo con l'ambiente fisico esterno. Ed è così che la nuova icona della mente è diventato il robot.

3. Robot futuri

La robotica oggi è un settore di ricerca molto attivo, specialmente per le applicazioni pratiche dei robot. Anzi le applicazioni sono così

preponderanti che comprimono le possibilità che la robotica possiede come strumento di ricerca puramente conoscitiva. Le applicazioni spingono verso direzioni che magari non sono quelle più interessanti dal punto di vista puramente scientifico, distolgono il ricercatore dall'esaminare con sistematicità se e quanto un robot si comporta come l'organismo naturale che si vorrebbe studiare e che il robot in qualche modo simula, e spesso creano una situazione che non né ricerca di base né ricerca applicata - una situazione che, per quanto possa avere anche dei vantaggi, finisce per rendere poco chiari i criteri stessi con cui giudicare quello che si sta facendo.

In ogni caso, specialmente se la robotica deve aiutarci a capire meglio il comportamento, le capacità, e la vita psichica di quei particolari organismi che sono gli esseri umani, il robot del futuro dovranno essere parecchio diversi da quelli di oggi. Oggi esistono i cosiddetti robot umanoidi, tipo il robot Asimo, ma si tratta di una somiglianza molto superficiale: una forma vagamente umana, camminare in modo bipede, produrre movimenti della faccia e del corpo che dovrebbero comunicare emozioni, usare il linguaggio con tutti i limiti molto seri degli attuali sistemi di riconoscimento e di sintesi automatica ereditati dall'intelligenza artificiale. I robot futuri dovranno invece muoversi in nuove direzioni di ricerca. Qui ne elenco cinque.

Prima di tutto, se vogliamo che i robot un giorno abbiano effettivamente sentimenti e emozioni, cioè "sentano" sentimenti e emozioni e non si limitino a "fare le facce" come i robot attuali, è necessario sviluppare una robotica interna che oggi non c'è. Che cosa è una robotica interna? Il corpo di un organismo contiene al suo interno non solo il cervello ma anche tutta una serie di altre strutture, organi e sistemi: cuore, intestino, sistema endocrino, sistema immunitario, e così via. Il cervello non interagisce soltanto con l'ambiente esterno, con quello che sta fuori del corpo, ma interagisce anche con questo ambiente interno, con quello che sta dentro al corpo, e il comportamento, specialmente nelle sue componenti dinamiche, motivazionali, emotive, più che in quelle strettamente cognitive, dipende dalle interazioni del cervello con gli organi e i sistemi interni del corpo. Il cervello riceve influenze da dentro al corpo, ad esempio influenze ormonali, e produce cambiamenti dentro al corpo, ad esempio contrazioni di muscoli interni e produzione di ormoni da parte del sistema endocrino. Le malattie neurologiche sono dovute a lesioni del sistema nervoso ma le malattie psichiatriche e i disturbi psicologici sono dovuti a malfunzionamenti di questi circuiti interni tra cervello e resto del corpo, e se la nuova scienza della mente deve spiegare queste malattie e questi disturbi, deve simulare gli organi e i sistemi interni del corpo e le loro interazioni con il cervello. Riprodurre in un artefatto costruito da noi questi organi e sistemi interni e queste interazioni è la robotica interna. La robotica interna è

una delle direzioni nuove in cui la robotica deve andare se i robot ci debbono aiutare a capire non solo gli aspetti cognitivi ma anche quelli dinamici e emotivi del comportamento.

La seconda direzione è la robotica mentale. Il cervello riceve input sensoriali dall'ambiente esterno, e anche da dentro al corpo, e risponde a questi input o con dei movimenti del corpo, i movimenti delle gambe, delle mani, degli occhi, degli organi fono-articolatori, o modificando lo stato interno del corpo. Ma negli animali che hanno un cervello più complesso, e soprattutto negli esseri umani, il cervello è capace anche di alimentarsi da solo. Normalmente quando abbiamo una esperienza, cioè quando nel sistema nervoso si realizza un particolare pattern di attivazione nei neuroni senso-motori, l'esperienza è dovuta all'arrivo di input sensoriali dall'ambiente esterno e ai movimenti con cui rispondiamo a questi input sensoriali. Ma gli esseri umani hanno anche una vita mentale. La vita mentale è la capacità di un cervello di avere esperienze auto-generate al suo stesso interno, cioè esperienze simili a quelle normali ma che sono dovute a input sensoriali che sono generati dentro allo stesso cervello e non arrivano da fuori, e a movimenti che sono programmati dentro al cervello ma sono eseguiti fisicamente e non si traducono in movimenti effettivi dei muscoli. E' in questo modo che abbiamo immagini mentali, ricordi, pensieri, sogni, allucinazioni, facciamo previsioni e piani, siamo consapevoli e coscienti. Gli animali più semplici non hanno queste cose, e sono sempre dipendenti dall'ambiente esterno, sempre in presa diretta con l'ambiente esterno. Gli esseri umani sono anche loro in presa diretta con l'ambiente esterno ma sono anche capaci di una vita mentale che li rende autonomi dall'ambiente esterno. Riprodurre la vita mentale in un artefatto costruito da noi è il compito della robotica mentale, che è la seconda direzione in cui la robotica futura dovrà andare.

La terza direzione è quella della robotica sociale. Gli organismi vivono e interagiscono con un ambiente che, oltre a varie altre cose, contiene anche altri organismi della loro stessa specie. Le interazioni con gli altri organismi della propria stessa specie costituiscono la vita sociale. Ci sono specie che hanno una vita sociale molto intensa e complessa, e altre specie che sono meno sociali. Gli esseri umani sono molto sociali. Praticamente fanno tutto insieme agli altri e interagendo e collaborando con gli altri. Oggi si comincia a studiare la socialità con i robot ma si tratta di una socialità elementare, più da insetti che da esseri umani. La robotica deve fare ancora molta strada se vuole riprodurre la socialità umana e vuole creare società umane artificiali. Nelle società animali gli individui all'interno di un gruppo sono quasi sempre parenti tra loro, cioè hanno gli stessi geni. Avere gli stessi geni significa si traduce nel fatto che uno si sacrifica per l'altro dato che in questo modo sacrifica sé stesso come individuo ma avvantaggia i propri geni, nella copia di

tali geni che è in possesso dell'altro individuo. Invece gli esseri umani formano società con individui che sono degli estranei, cioè individui che non hanno gli stessi geni. Questo crea innumerevoli problemi perché ogni individuo tende a pensare a sé stesso ma nello stesso tempo deve cooperare con gli altri perché cooperando con gli altri si possono ottenere cose che da soli non si possono ottenere: si possono produrre beni in modo organizzato, si possono effettuare scambi di beni, si può modificare l'ambiente in modo da migliorarlo, si possono creare meccanismi sociali che impediscono agli individui di essere troppo egoisti. Il contrasto tra queste due tendenze costituisce una fonte costante di tensioni, e tutta la vita sociale e politica serve in sostanza a gestire queste tensioni. Riprodurre tutto questo in società di robot è il compito della robotica sociale, di cui oggi si vedono appena gli inizi.

Un aspetto centrale della vita sociale umana è il linguaggio e il fatto che quasi tutte le interazioni tra gli esseri umani passano attraverso il linguaggio. Il linguaggio umano è qualcosa di molto complicato e, nonostante che ci siano tante scienze che lo studiano, la scienza lo capisce ancora poco. La robotica è essenziale per lo studio del linguaggio. Come pensava Wittgenstein, il significato di una parola è il ruolo che la parola svolge in quelli che lui chiamava i "giochi linguistici", cioè le specifiche situazioni, con le loro specifiche regole, all'interno delle quali la parola viene usata, pronunciata da uno e capita da un altro. Il significato di una parola è quello che uno fa con la parola. Questa visione operazionistica del significato delle parole si può studiare bene soltanto con i robot. I robot vivono in un ambiente e interagiscono con l'ambiente, che nella robotica sociale include altri robot. Il significato delle parole per loro, come per noi, sono le specifiche situazioni, con le loro specifiche anche se implicite regole, in cui ciascuna parola viene usata. Rispetto a Wittgenstein, i vantaggi sono due. Il primo è che le situazioni in cui le parole sono usate non vengono più soltanto immaginate e descritte a parole, come facevano i filosofi wittgensteiniani, ma sono situazioni fisiche che si possono osservare oggettivamente e manipolare sperimentalmente. Il secondo vantaggio è che i robot hanno un cervello e il significato delle parole può essere, letteralmente, osservato guardando che cosa succede nel loro cervello artificiale quando usano il linguaggio, un po' come facciamo con i cervelli reali usando le neuroimmagini. La robotica linguistica è la quarta delle direzioni future della robotica, e solo con la robotica linguistica potranno essere superati i limiti così evidenti degli attuali sistemi di comprensione e di produzione automatica del linguaggio.

La quinta e ultima direzione è la robotica culturale. Un altro aspetto fondamentale del comportamento umano è che la maggior parte di quello che un essere umano sa e sa fare non lo eredita geneticamente e non lo apprende interagendo con la natura, come

succede per gli animali, ma lo apprende da altri esseri umani o, come si dice, culturalmente. Si tratta allora prima di tutto di costruire robot che imparano dagli altri, cioè che sappiano confrontare il proprio comportamento con il comportamento di un robot modello e sappiano modificare il proprio comportamento per renderlo più simile a quello del modello. L'altro robot può non limitarsi a funzionare come modello in modo puramente passivo ma può darsi da fare perché l'allievo impari, e allora quello che era un robot modello diventa un robot insegnante. Un robot insegnante deve saper richiamare l'attenzione dell'allievo su una certa particolare cosa, deve saper descrivere a parole il comportamento da apprendere invece che soltanto dimostrarlo concretamente, deve ripetere più volte le cose che per l'allievo sono più difficili da imparare, deve valutare i progressi dell'allievo, deve approvare o criticare il comportamento dell'allievo.

Poi la robotica culturale deve cercare di capire chi impara da chi e chi insegna a chi. Chi sono gli insegnanti e chi sono gli allievi? Gli individui della generazione precedente, gli adulti, insegnano agli individui della generazione successiva, ai giovani, oppure, come avviene sempre di più oggi, i giovani imparano soprattutto dagli altri giovani e non hanno molta voglia di imparare dagli adulti? E come si scelgono gli specifici individui da cui imparare? Ci si basa sul fatto che sono gli individui che hanno più successo, anche se non sappiamo se il loro successo è dovuto proprio al comportamento che impariamo da loro o ad altri fattori? E poi si impara non solo dalle persone ma anche dagli artefatti tecnologici. Se usiamo un certo artefatto tecnologico, finiamo per comportarci e per pensare nei modi previsti da quel particolare artefatto tecnologico. E' come se l'artefatto tecnologico trasferisse a noi i comportamenti e i modi di pensare che sono incorporati in esso. Come scegliamo gli artefatti tecnologici da cui imparare, cioè in pratica gli artefatti tecnologici da comprare e usare?

E poi si tratta di costruire culture di robot, come esistono le culture o le civiltà umane. Se due individui interagiscono, tendono a diventare simili. Se un insieme di individui formano una comunità di individui che interagiscono molto tra di loro mentre interagiscono poco con individui di altre comunità, si creano una serie di culture diverse e separate, cioè una serie di modi di fare le cose, di pensare, una serie di valori condivisi all'interno di una comunità ma diversi da quelli delle altre comunità. Sono queste le culture umane o, come si dice anche, per le culture considerate più importanti, le civiltà umane. La robotica culturale deve studiare tutte queste cose, fino ad arrivare a riprodurre la globalizzazione culturale che sta avvenendo oggi, cioè l'emergere di una cultura unica dovuta al fatto che le interazioni tra gli individui non sono più soltanto locali e dipendenti dal fatto di vivere vicini ma, con lo sviluppo delle tecnologie per trasportare gli

individui, i beni e le informazioni, diventano interazioni globali, cioè potenzialmente di ogni individuo con ogni altro individuo, indipendentemente da dove stanno nello spazio.